

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANTS: Aaron Kaplan et al.
- (ii) TITLE OF INVENTION: ENHANCING INORGANIC CARBON FIXATION BY PHOTOSYNTHETIC ORGANISMS
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
- (B) STREET: 2001 Jefferson Davis Highway, Suite 207
- (C) CITY: Arlington
- (D) STATE: Virginia
- (E) COUNTRY: United States of America
- (F) ZIP: 22202
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
- (B) COMPUTER: Twinhead Slimnote-890TX
- (C) OPERATING SYSTEM: MS DOS version 6.2,
Windows version 3.11
- (D) SOFTWARE: Word for Windows version 2.0 converted to
an ASCII file
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Friedmam, Mark M.
- (B) REGISTRATION NUMBER: 33,883
- (C) REFERENCE/DOCKET NUMBER: 325/45
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 972-3-5625553
- (B) TELEFAX: 972-3-5625554
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4957
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

AAGCTTGGAT TGAAGCGATC GGGGTCAATC CCAGCGATGA TCCTCAGTTC   50
CTCCTGATGG TCATCCCTT TAGCGCCAAG ATTGAGGATC TGCTGCAAGG  100
GCTGGATTTC GCCTATCCCG AGGCCGTGAA AGTGGGCGGA TTGGCCAGTG  150
GTTTGGGGGC AGAGTCAGCG ATCGCCAGCT TGTTTTTCA AGACCGACAG  200
GTCGATGGCG TGATTGGGCT AGCCCTCAGT GGCAATGTCC AGCTGCAGGC  250
GATCGTGGCT CAGGGCTGTC GTCCAGTTGG CCCGCTTTGG CATGTGGCAG  300
CGGCGGAGCG CAACATTCTG CGGCAACTTC AGACCGAAGA CGAGGAACCG  350
ATCGCCGCGC TGAAGGCCCT ACAGTCAGTC CTGCGTGATC TCTCCCTGA  400
ATTACAGCGA TCCTCTGTG TGGGCCTGGC CTGCAATTCT TTCCAAACGG  450
TATTACAACC GGGCGACTTC CTGATCCGTA ACCTGCTGGG GTTTGATCCC  500
CGCACTGGTG CTGTAGCAAT CGGCGATCGC ATTCGAGTTG GGCAGCGGCT  550
GCAGCTGCAC GTACGGGATG CCCAGACAGC GGCGGATGAC CTCGAGCGGC  600
AACTGGGGCA ATGGTGCCGG CAGCATGCGA CAAAACCAGC AGCTTCCCTC  650
TTGTTTTCCT GCTTGGGGCG CGGCAAGCCC TTCTATCAGC AGGCCAACTT  700

```

CTAGATCGCAA	CTGATTTCAGC	ATTACCTCTC	AGAGCTGCCC	CTAGCTGGCT	750
TTTTCTGTAA	TGGCGAAATC	GGCCCCGATCG	CTGGCAGCAC	CTACCTGCAT	800
GGCTACACAT	CGGTGCTGGC	TTTGCTGTCTG	GCCAAAACCTC	ACTAGCGCCA	850
GCGAGACCTG	ATTGTCGATC	TGCTGAGCGC	GA CTGTAGCG	CTGGAAATAG	900
GCCCCGACCT	GAGCAGGCGC	ATCGGCCAAG	CTGACCGTAG	TATCACCGTC	950
AGCCACCCCC	GCCCCAGAAAT	TCCGCAACAT	CGGCAGGAGA	GCGATCCGCT	1000
CCGCCTCCGA	TAAATTCAAC	GGCTCATGGG	TCAACAGGCG	GATCAAGTAC	1050
TCTGACTGCG	ATCGCCATCC	ATTCGCCGCC	AAAACGTTTG	TAAATCAGTC	1100
TTGATCCGGT	AGCGATCGCA	CCC GACGGGA	CTCTAGTTCT	AGTTGCCAAC	1150
CTTCAGCGGC	AGGTTGTACG	GTTCCGAGTC	GGTAGGGATG	GGGATAGCTG	1200
ACCAAGGAAC	CGGTCGTGAC	TTCCACAGAA	GCACCTTGCT	GACTGGTGCG	1250
TTGAGTGTGG	AGGTTGGCTG	TGAAGATCAC	CGAGACCGTG	CCCGCTTCGA	1300
GGATTGATCG	CAATTCTCTCG	GCAATTTTCTA	AGATGTAGCG	CTGACCAAGC	1350
GGATGCTGCT	GTTGATCGGG	CAGATGCTCC	AACACATTGT	GGTGAATCAT	1400
CACCCAGCGT	TGGCTAGCGG	TGGAAGTGCG	GAGTTCCTGT	TGCAGCCAGT	1450
TGAGTTGCGC	GCAATCGACT	CGCCCCGAT	GCAGTTGATG	GCCCGCTTCA	1500
TCAAAAGCGA	TCGAATTTCAG	CGCAAACAGA	TCGAGATCCG	GTGCGATCGT	1550
GCAGCGATAG	TAGGGGGCGAT	CGCTCGTGAA	GCCAAAGTCT	TGATAGAGCT	1600
CGACAAACTC	GCCACACCGG	GTGCGATCGC	GATCGCTCGC	TGCGGCGGGC	1650
ATATCTGTGT	TGCCCGGCGC	CACATAGACC	GGATAGGGCA	ACTGGCGCAA	1700
TTGTTGCAGC	AGCCACTGAT	GGTTTTCCCG	CTCCCCGTGC	TGGGTTAAAT	1750
CCCCCGGCAG	CAACAGGAAG	TCCAAATCCA	GCCTGCCAG	TTCTGTCAAG	1800
ATTTGCTCAA	AAGCCGGAAT	GCTGCACTCA	ATCAAATGGA	AGCGATGGGG	1850
ATGGTGCCAA	ATTGTCGTGCG	GCAGTCCAAT	GTGGAGATCG	CTCAGCAGCG	1900
CAAATCGAAA	CGCTCGGTTT	ATTGCCATCC	CCTCAGCTAT	CGAGCCCGAT	1950
TCTAGGCGAA	GCTAGGTGCA	GTCCGTTGTC	TTCACTTGCA	AGCAATTCATG	2000
CGCAGAGTTT	CGGTTTCGCA	CACGCTCAAT	CCGCTCTCTC	AGAAATTTCCA	2050
AGTGGTTCAG	ACTTTGCCCG	ATGCGCAACA	GGTCTATGCG	GACTGCCATC	2100
GCCCGCTGCA	TTTGATATTT	GGCTGTGCTC	GCGGGCGCTT	TCTGCTGGCA	2150
ATGGCGACAC	GACAACCTGA	GTGGAATTAT	CTGGGGCTGG	AAATTCGTGA	2200
GCCGCTGGTA	GATGAGGCGA	ACGCGATCGC	CCGCGAACGT	GAAGTACCA	2250
ATCTCTACTA	CCACTTCAGC	AACGCCAATT	TGGACTTGGA	ACCGCTGCTG	2300
CGATCGCTGC	CGACAGGGAT	TTTGACGCGG	GTGAGCATTC	AGTTCCCGGA	2350
TCCTTGCTTC	AAGAAACGCC	ATCAAAGCG	ACGCGTCGTC	CAGCCGGAAC	2400
TGGTGAAGC	CCTCGCGACT	GCGTTACCTG	CTGGTGCAAG	GGTCTTTCTG	2450
CAATCCGATG	TGCTGGAAGT	GCAGGCAGAG	ATGTGCGAAC	ACTTTGCGGC	2500
GGAACCCCGC	TTTCAGCGCA	CCTGCTTGGA	CTGGCTGCCG	GAATATCCGC	2550
TGCCCCGTCC	GACCGAGCGC	GAAATTGCCG	TTCAAAACAA	ACAGTTGCCA	2600
GTCTACCGTG	CTCTCTTCAT	TCGGCAGCCA	GCGGACTAAG	CTCTTAAGGC	2650
AAGCGTTAGC	GCGATCGCGA	TGACTGTCTG	GCAAACTCTG	ACTTTTGCCC	2700
ATTACCAACC	CCAACAGTGG	GGCCACAGCA	GTTTCTTGCA	TCTGGCTGTTT	2750
GGCAGCTGCG	GAGCTTGGCG	GGCCTCCAGC	GACCTGTTGG	TTTGGTCTGA	2800
GGCACTGGGT	GGCTTCTTGC	TTGCTGTCTG	CTACGGTTCC	GCTCCGTTTG	2850
TGCCCCAGTT	CGCCCCAGGG	TTGGGGCTAG	CCGCGATCGC	GGCCTATTGG	2900
GCCCTGCTCT	CGCTGACAGA	TATCGATCTG	CGGCAAGCAA	CCCCCATTTA	2950
CTGGCTGGTG	CTGCTCTACT	GGGGCGTCTG	TGCCCTAGCA	ACGGGACTCT	3000
CACCCGTACG	CGCTGCAGCT	TTAGTTGGGC	TAGCCAAACT	GACGCTCTAC	3050
CTGTTGGTTT	TTGCCCTAGC	GGCTCGGGTT	CTCCGCAATC	CCGCTCTGCG	3100
ATCGCTGCTG	TTCTCGGTCG	TCGTGATCAC	ATCGCTTTTT	GTCAGTGTCT	3150
ACGGCCCTCA	CCAATGGATC	TACGGCGTTG	AAGAGCTGGC	GACTTGGGTG	3200
GATCGCAACT	CGGTTGCCGA	CTTCACCTCA	CGGGTTTACA	GCTATCTGGG	3250
CAACCCCAAC	CTGCTGGCTG	CTTATCTGGT	GCCGACGACT	GCCTTTTCTG	3300
CAGCAGCGAT	CGGGGTGTGG	CGCGGCTGGC	TCCCCAAGCT	GCTGGCGATC	3350
GCTCGACAG	GCTCGAGCAG	CTTATGCTCT	ATCCTCACCT	ACAGTCGGCG	3400
TGGCTGGCTG	GGTTTGTCTG	CCATAGCTTTT	TGCTGGGGCG	TTATTAGGGC	3450
TCTACTGGTT	TCAACCCCGT	CTACCCGCAC	CTGTGGCGAC	CTGGCTATTG	3500
CCAGTCGTAT	TGGGTGGACT	AGTCGCGGGT	CTCTGGGTGG	CGGTGCTTGG	3550

ACTTGAGCCG TTGCGCGTGC GCGTGTGAG CATCTTTGTG GGGCGTGAAG 3600
 ACAGCAGCAA CAACTTCCGG ATCAATGTCT GGCTGGCGGT GCTGCAGATG 3650
 ATTCAAGATC GGCCTTGGCT GGGCATCGGC CCCGGCAATA CCGCCTTAA 3700
 CCTGGTTTAT CCCCTCTATC AACAGGCGCG CTTTACGGCG TTGAGCGCCT 3750
 ACTCCGTCCC GCTGGAAGTC GCGGTTGAGG GCGGACTACT GGGCTTGACG 3800
 GCCTTCGCTT GGCTGCTGCT GGTACGCGG GTGACGCGG TCGGCGAGGT 3850
 GAGCCGACTG CGGCGCGATC GCAATCCCCA AGCCTTTTGG TTGATGGCTA 3900
 GCTTGGCCGG TTTGGCAGGA ATGCTGGGTC ACGGTCTGTT TGATACCGTG 3950
 CTCTATCGAC CGGAAGCCAG TACGCTCTGG TGGCTCTGTA TTGGAGCGAT 4000
 CGCGAGTTTC TGGCAGCCCC AACCTTCCAA GCAACTCCCT CCAGAAGCCG 4050
 AGCATTGAGA CGAAAAATG TAGCGGGGTC CCCAACAAAT TCCTGTGCAC 4100
 CCGACTGGAT CCACCACCTA AACTGGATCC CAAAGGTATC CGGTGGATCT 4150
 AGGGTCATAA CGAACTCCGA CCGCGATCGC GTCCGCGAAC TGAACCTCCA 4200
 TCGCACCGAA GCGGAGTTCG TTAGTCGTTG AAGAGCCAAT GCTAGAGGGG 4250
 GCTGCCAAG CAGTTGGGCT GGAAGCAGGC TCGGAGAAGC CACCCGCATC 4300
 CAAGGCAAAG TTCAGCCGAC CTTCCGAAA GACTACGATC GCCACGGCGG 4350
 CTCTGCCAGC TAAGTCAGCG CTGGGTTAGT TGTCATAGCA GTCCGAGATC 4400
 AAGTTAGGAC AACTTCATAG AGGGACTCGC TCAGAGTCAA CAGCCGCTGT 4450
 CCGTGGGGGT GCGCAATCAC CCCACACCC ACGCACTGGG GGAAGTGGAT 4500
 CCCCCAGGCC CCCCCAACA AGATTTCGGA TAAGGGGCAT CGGCTGAATC 4550
 GCGATCGCTG CGGGTAAAAC TAGCCGGTGT TAGCCATGGG TTTGAGACTA 4600
 ATCGGCACGG GGCAAAACGT CCTGATTAT TGTCTCAATG TGATAGGTTA 4650
 CATCGTCAAA AACAAGGCC AAGAGGTAGG AAAAATCACG ACCGCCCAAG 4700
 TCCGAGGGCT TTGCTGTTGG GAGCGACCTA GGGCAGACTA GACAGAGCAT 4750
 TGCTGTGAGC CAAAGCGCCT TCAATTGCTG GCGGCTGTGG GTTTTTCGGA 4800
 GGTTCGAAA TGAAAGACCT TTTCTGCAAT GTCTCCGCT ATCCCGCTA 4850
 CTTTCATACC TTCCAGCTGG GTATTTTTTA GTCGATCTAC CAGTGGGTGC 4900
 GGCCGATGGT TCGCAACCCA GTCGCGGCTT GGGCGCTGCT AGGCTTTGGA 4950
 GTTTCGA 4957

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1404
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGACTGTCT GGCAAACTCT GACTTTTGCC CATTACCAAC CCCAACAGTG 50
 GGGCCACAGC AGTTTCTTGC ATCGGCTGTT TGGCAGCCTG CGAGCTTGGC 100
 GGGCCTCCAG CCAGCTGTTG GTTTGGTCTG AGGCACTGGG TGGCTTCTTG 150
 CTTGCTGTGC TCTACGGTTC GGCTCCGTTT GTGCCAGTT CCGCCCTAGG 200
 GTTGGGGCTA GCGCGATCG CGGCCTATTG GGCCTGCTC TCGCTGACAG 250
 ATATCGATCT GCGCAAGCA ACCCCATTC ACTGGCTGGT GCTGCTCTAC 300
 TGGGGCGTCG ATGCCCTAGC AACGGGACTC TCACCCGTAC GCGCTGCAGC 350
 TTTAGTTGGG CTAGCCAAAC TGACGCTCTA CCTGTTGGTT TTTGCCCTAG 400
 CGGCTCGGGT TCTCCGCAAT CCCCGTCTGC GATCGCTGCT GTTCTCGGTC 450
 GTCGTGATCA CATCGCTTTT TGTCAGTGTC TACGGCCTCA ACCAATGGAT 500
 CTACGGCGTT GAAGAGCTGG CGACTTGGGT GGATCGCAAC TCGGTTGCCG 550
 ACTTCACCTC ACGGGTTTAC AGCTATCTGG GCAACCCCAA CCTGTGGGCT 600
 GCTTATCTGG TGCCGACGAC TGCCCTTTTCT GCAGCAGCGA TCGGGGTGTG 650
 GCGCGGCTGG CTCCCCAAGC TGCTGGCGAT CGCTGCGACA GGTGCGAGCA 700
 GCTTATGTCT GATCCTCACC TACAGTCGCG GTGGCTGGCT GGGTTTGTG 750
 GCCATGATTT TTGCTGGGC GTTATTAGGG CTCTACTGGT TTCAACCCCG 800
 TCTACCCGCA CCCTGGCGAC GCTGGCTATT CCCAGTCGTA TTGGGTGGAC 850
 TAGTCGCGGT GCTCTTGGTG GCGGTGCTTG GACTTGAGCC GTTGC GCGTG 900
 CGCGTGTTGA GCATCTTTGT GGGGCGTGAA GACAGCAGCA ACAACTTCCG 950
 GATCAATGTC TGGCTGGCGG TGCTGCAGAT GATTCAAGAT CGGCCTTGGC 1000

(2) INFORMATION FOR SEO ID NO:3:

(A) LENGTH: 467
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

Met	Thr	Val	Trp	Gln	Thr	Leu	Thr	Phe	Ala	His	Tyr	Gln	Pro	Gln
				5					10					15
Gln	Trp	Gly	His	Ser	Ser	Phe	Leu	His	Arg	Leu	Phe	Gly	Ser	Leu
				20					25					30
Arg	Ala	Trp	Arg	Ala	Ser	Ser	Gln	Leu	Leu	Val	Trp	Ser	Glu	Ala
				35					40					45
Leu	Gly	Gly	Phe	Leu	Leu	Ala	Val	Val	Tyr	Gly	Ser	Ala	Pro	Phe
				50					55					60
Val	Pro	Ser	Ser	Ala	Leu	Gly	Leu	Gly	Leu	Ala	Ala	Ile	Ala	Ala
				65					70					75
Tyr	Trp	Ala	Leu	Leu	Ser	Leu	Thr	Asp	Ile	Asp	Leu	Arg	Gln	Ala
				80					85					90
Thr	Pro	Ile	His	Trp	Leu	Val	Leu	Leu	Tyr	Trp	Gly	Val	Asp	Ala
				95					100					105
Leu	Ala	Thr	Gly	Leu	Ser	Pro	Val	Arg	Ala	Ala	Ala	Leu	Val	Gly
				110					115					120
Leu	Ala	Lys	Leu	Thr	Leu	Tyr	Leu	Leu	Val	Phe	Ala	Leu	Ala	Ala
				125					130					135
Arg	Val	Leu	Arg	Asn	Pro	Arg	Leu	Arg	Ser	Leu	Leu	Phe	Ser	Val
				140					145					150
Val	Val	Ile	Thr	Ser	Leu	Phe	Val	Ser	Val	Tyr	Gly	Leu	Asn	Gln
				155					160					165
Trp	Ile	Tyr	Gly	Val	Glu	Glu	Leu	Ala	Thr	Trp	Val	Asp	Arg	Asn
				170					175					180
Ser	Val	Ala	Asp	Phe	Thr	Ser	Arg	Val	Tyr	Ser	Tyr	Leu	Gly	Asn
				185					190					195
Pro	Asn	Leu	Leu	Ala	Ala	Tyr	Leu	Val	Pro	Thr	Thr	Ala	Phe	Ser
				200					205					210
Ala	Ala	Ala	Ile	Gly	Val	Trp	Arg	Gly	Trp	Leu	Pro	Lys	Leu	Leu
				215					220					225
Ala	Ile	Ala	Ala	Thr	Gly	Ala	Ser	Ser	Leu	Cys	Leu	Ile	Leu	Thr
				230					235					240
Tyr	Ser	Arg	Gly	Gly	Trp	Leu	Gly	Phe	Val	Ala	Met	Ile	Phe	Val
				245					250					255
Trp	Ala	Leu	Leu	Gly	Leu	Tyr	Trp	Phe	Gln	Pro	Arg	Leu	Pro	Ala
				260					265					270
Pro	Trp	Arg	Arg	Trp	Leu	Phe	Pro	Val	Val	Leu	Gly	Gly	Leu	Val
				275					280					285
Ala	Val	Leu	Leu	Val	Ala	Val	Leu	Gly	Leu	Glu	Pro	Leu	Arg	Val
				290					295					300

(2) INFORMATION FOR SEQ ID NO:4:

(A) LENGTH: 1425
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

ATGGTGTCTC	CCATCTCTAT	CTGGCGATCG	CTGATGTTTG	CGGGTTTTTC	50
CCCCCAGGAA	TGGGGCCGGG	GCAGTGTGCT	CCATCGTTTG	GTGGGCTGGG	100
GACAGAGTTG	GATACAGGCT	AGTGTGCTCT	GGCCCCACTT	CGAGGCATTG	150
GGTACGGCTC	TAGTGGCAAT	AATTTTTATT	GCGGCTCCCT	TCACCTCCAC	200
CACCATGTTG	GGCATPTTTA	TGCTGCTCTG	TGGAGCCTTT	TGGGCTCTGC	250
TGACCTTTGC	TGATCAACCA	GGGAAGGGTT	TGACTCCCAT	CCATGTTTTA	300
GTTTTTGCCT	ACTGGTGCAT	TTCGGCGATC	GCCGTGGGAT	TTTCTCCGGT	350
AAAAATGGCG	CGCGCGTCGG	GGTTAGCGAA	ATTAACAGCT	AATTTATGTC	400
TGTTTCTACT	GGCGGCGAGG	TTATTGCAA	ACAAACAATG	GTTGAACCGG	450
TTAGTAACCG	TTGTTTTACT	GTAGAGGCTA	TTGGTGGGGA	GTTACGGTCT	500
GGCACAACAG	TGGGACGGGG	TAGAACAGTT	AGGCACTTGG	AAGTACCCCA	550
CCTCTACCTT	GGCCCAGGCC	ACTAGGGTAT	ATAGCTTTTT	GGGTAATCCC	600
AATCTCTTGG	CGGCTTACCT	GGTGCCCATG	ACGGGTTTGA	GCTTGAGTGC	650
CCTGGTGGTA	TGGCGACGGT	GGTGGCCCAA	ACTGCTGGGA	GCAACCATGG	700
TGATTGTTAA	CCTACTCTGT	CTCTTTTTTA	CCCAGAGCCG	GGGCGGTTGG	750
CTAGCAGTGC	TGGCCCTGGG	AGCTACCTTC	CTGGCCCTTT	GTTACTTCTG	800
GTGGTTACCC	CAATTACCCA	AATTTTGCA	ACGGTGGTCT	TTGCCCTTGG	850
CGATCGCCGT	GGCGGTTATA	TTAGGTGGGG	GAGCGTTGAT	TGCGGTGGAA	900
CCGATTTCGAC	TCAGGGCCAT	GAGCATTTTT	GCTGGGCGGG	AAGACAGCAG	950
TAATAATTTT	CGCATCAATG	TTTGGGAAGG	GGTAAAAGCC	ATGATCCGAG	1000
CCCGCCCTACT	CATTGGCATT	GGCCCAGGTA	ACGAAGCCTT	TAACCAAATT	1050
TATCCTTACT	ATATGCGGGC	CCGCTTCACC	GCCCTGAGTG	CCTATTCCAT	1100
TTACCTAGAA	ATTTTGGTGG	AAACGGGTGT	AGTTGGTTTT	ACCTGTATGC	1150
TCTGGCTGTT	GGCCGTTACC	CTAGGCAAAG	CGGTAGACTT	GGCTAAACGC	1200
TGTCGCCAAA	CCCTCGCCCC	GGGAAGGCATC	TGGATTATGG	GGGCTTTAGC	1250
GGCGATCATC	GGTTTGTTGG	TCCACGGCAT	GGTAGATACA	GTCTGGTACC	1300

GTCCCCCGGT GAGCACTTTG TGGTGGTTGC TAGTGGCCAT TGTTGCTAGT 1350
 CAGTGGGCCA GCGCCAGGC CCGTTTGGAG GCCAGTAAAG AAGAAAATGA 1400
 GGACAAACCT CTTCTTGCTT CATAA 1425

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 474
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Val Ser Pro Ile Ser Ile Trp Arg Ser Leu Met Phe Gly Gly		
	5	10
		15
Phe Ser Pro Gln Glu Trp Gly Arg Gly Ser Val Leu His Arg Leu		
	20	25
		30
Val Gly Trp Gly Gln Ser Trp Ile Gln Ala Ser Val Leu Trp Pro		
	35	40
		45
His Phe Glu Ala Leu Gly Thr Ala Leu Val Ala Ile Ile Phe Ile		
	50	55
		60
Ala Ala Pro Phe Thr Ser Thr Thr Met Leu Gly Ile Phe Met Leu		
	65	70
		75
Leu Cys Gly Ala Phe Trp Ala Leu Leu Thr Phe Ala Asp Gln Pro		
	80	85
		90
Gly Lys Gly Leu Thr Pro Ile His Val Leu Val Phe Ala Tyr Trp		
	95	100
		105
Cys Ile Ser Ala Ile Ala Val Gly Phe Ser Pro Val Lys Met Ala		
	110	115
		120
Ala Ala Ser Gly Leu Ala Lys Leu Thr Ala Asn Leu Cys Leu Phe		
	125	130
		135
Leu Leu Ala Ala Arg Leu Leu Gln Asn Lys Gln Trp Leu Asn Arg		
	140	145
		150
Leu Val Thr Val Val Leu Leu Val Gly Leu Leu Val Gly Ser Tyr		
	155	160
		165
Gly Leu Arg Gln Gln Val Asp Gly Val Glu Gln Leu Ala Thr Trp		
	170	175
		180
Asn Asp Pro Thr Ser Thr Leu Ala Gln Ala Thr Arg Val Tyr Ser		
	185	190
		195
Phe Leu Gly Asn Pro Asn Leu Leu Ala Ala Tyr Leu Val Pro Met		
	200	205
		210
Thr Gly Leu Ser Leu Ser Ala Leu Val Val Trp Arg Arg Trp Trp		
	215	220
		225
Pro Lys Leu Leu Gly Ala Thr Met Val Ile Val Asn Leu Leu Cys		
	230	235
		240
Leu Phe Phe Thr Gln Ser Arg Gly Gly Trp Leu Ala Val Leu Ala		
	245	250
		255
Leu Gly Ala Thr Phe Leu Ala Leu Cys Tyr Phe Trp Trp Leu Pro		
	260	265
		270
Gln Leu Pro Lys Phe Trp Gln Arg Trp Ser Leu Pro Leu Ala Ile		
	275	280
		285
Ala Val Ala Val Ile Leu Gly Gly Gly Ala Leu Ile Ala Val Glu		
	290	295
		300
Pro Ile Arg Leu Arg Ala Met Ser Ile Phe Ala Gly Arg Glu Asp		
	305	310
		315
Ser Ser Asn Asn Phe Arg Ile Asn Val Trp Glu Gly Val Lys Ala		
	320	325
		330
Met Ile Arg Ala Arg Pro Ile Ile Gly Ile Gly Pro Gly Asn Glu		
	335	340
		345

Ala Phe Asn Gln Ile Tyr Pro Tyr Tyr Met Arg Pro Arg Phe Thr		
350	355	360
Ala Leu Ser Ala Tyr Ser Ile Tyr Leu Glu Ile Leu Val Glu Thr		
365	370	375
Gly Val Val Gly Phe Thr Cys Met Leu Trp Leu Leu Ala Val Thr		
380	385	390
Leu Gly Lys Gly Val Glu Leu Val Lys Arg Cys Arg Gln Thr Leu		
395	400	405
Ala Pro Glu Gly Ile Trp Ile Met Gly Ala Leu Ala Ala Ile Ile		
410	415	420
Gly Leu Leu Val His Gly Met Val Asp Thr Val Trp Tyr Arg Pro		
425	430	435
Pro Val Ser Thr Leu Trp Trp Leu Leu Val Ala Ile Val Ala Ser		
440	445	450
Gln Trp Ala Ser Ala Gln Ala Arg Leu Glu Ala Ser Lys Glu Glu		
455	460	465
Asn Glu Asp Lys Pro Leu Leu Ala Ser		
470		

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGCTAGCCG CGATCGCGGC CTATTGGGCC C 31

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGCTAGGGA TCGCGCTAT TGGGCC 27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGCTCAGAT CGCGCCTATT GGGCCC 26

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Leu Ala Ala Ile Ala Ala Tyr Trp Ala Leu